

גלעד עיני 034744920

מתן פינץ 300895315

למידה ממוכנת

תרגיל בית מספר 3

Pca, Mda and NN

FisherFaces.m א.

- צעד נרמול (פונקציה מצורפת): לכל עמודה: החסרנו את הממוצע וחילקנו ב STD
- חישבנו את $AVZ = Wpca$
 - A זה הדאטה המנורמל
 - $A'A = S$
 - $[V, Z] = \text{eigs}(\text{Stag}, N-c)$ – להשיג את ה $N-c$ העליונים.
 - ירדנו ל $N-c$ מימדים
- חישבנו את $Wmda$
 - חישוב SW ו SB
 - כפלו כל אחת משני הצדדים ב $WPCA$
 - פתרנו שוב את בעיית הע"ע הכללית
 - $[Vm, \sim] = \text{eigs}(SbNew, SwNew, c-1)$
 - ירדנו ל $c-1$ מימדים
- $Wmda * Wpca' = Wopt$
- שמרנו את F (WOPT) ואת MN שזה ממוצע הדאטה

recogTest.m ב.

- צעד נרמול ל $TRAIN$
- חישבנו את המודל לכל חמישים האנשים על $TRAIN$
- הטלנו את המודלים על $WOPT$ (צמצמנו את המימדים)
- צעד נרמול ל $TEST$
- הטלנו את $TEST$ על $WOPT$ (צמצמנו את המימדים)
- שלחנו ל KNN וספרנו סיווגים נכונים

```
The recognition rate is 0.88
0.8800
```

Regression

a. כתבו פונקציה שתחשב מודל ותחזיר את W ואת שגיאת ה MSE על ה $TRAIN$ ועל ה $TEST$

- על ה $TRAIN$

- פיצלנו ל samples and labels
- חישבנו את w האופטימלי מההרצאה:

$$\hat{w} = (X^t X)^{-1} X^t y$$

- עכשיו שיש W הרצנו חישוב שגיאה על ה $TRAIN$ וגם על ה $TEST$

- הרצנו לולאה על כל הדוגמאות

- לוקחים את המרחק הריבועי בין Y האמיתי

$$f(x, \hat{w}) = \hat{w}_0 + \sum_{i=1} \hat{w}_i x_i,$$

- לבין $F(X, W)$ שהיא:

b. חילקנו ל $TRAIN$ ו $TEST$ תוך שימוש בפונקציה המצורפת permAndSplit

```
case split is 10 and 382
  On |trainData| =10
  The MSE:4.06. The of
  On |testData| =382
  The MSE:58.79
case split is 50 and 342
  On |trainData| =50
  The MSE:12.094. The
  On |testData| =342
  The MSE:18.6622
case split is 100 and 292
  On |trainData| =100
  The MSE:11.2385. The
  On |testData| =292
  The MSE:11.992
case split is 200 and 192
  On |trainData| =200
  The MSE:13.8375. The
  On |testData| =192
  The MSE:9.6525
```

c. הסבר על התוצאות:

- ניתן לראות שהכנו את הקו על 10 דוגמאות אימון, המודל נתפר במיוחד עבורם, ולא הצליח להכליל. קיבלנו 4 שגיאה באימון ו59 בטסט.
- ככל שמוסיפים נקודות אימון, הרגרשן צריך להתחשב בעוד נקודות ולכן השגיאה שלו גדלה, אבל הוא מצליח להכליל יותר טוב שזה בדיוק המטרה שלנו

מספר דוגמאות אימון	MSE TRAIN	MSE TEST
10	4	58
50	12	18
100	11	12
200	13	10

רואים שיצאו לנו פונקציות כמעט מונוטוניות (כמעט: בגלל הרנדום). $MSE TRAIN$ יותר גדולה שיש יותר דוגמאות לאימון, וה $MSE TEST$ של ההכללה משתפרת ככל שיש יותר דוגמאות אימון.

SVM אחד נגד כולם תוך שימוש בקרוס וואלידיישן (CV) למציאת הפרמטרים

א. `:trainLinearSVM.m`

- הרצנו CV לכל קלאס עם $C=1$ וגם $C=5$.
 - מיינו את ה-DATA כאשר הקלאס הנוכחי היה בראש ה-DATA הממויין (סמפלים ולייבלים בהתאמה).
 - הרצנו את SVMTRAIN עם שני ה-C וסכמנו את התוצאה
 - עבור $C=1$ קיבלנו `cv_acc = 95.55`
 - עבור $C=5$ קיבלנו `cv_acc = 95.69`
- אימנו 10 מסווגים ליניאריים עם $C=5$ ושמרנו אותם.

```
***
LinearSvm was trained with bestC = 5
***
Cross-validation accuracy is 95.69
```

ב. `testSVM.m` (על א):

- אתחלנו מטריצת הסתברויות
 - כל שורה: מייצגת תוצאות של טסט אחד
 - כל עמודה: מייצגת קלאס.
 - הערך במקום ה j , הוא הערך ההסתברותי שהטסט i שייך לקלאס ה j
- הרצנו לכל קלאס:
 - הפוך את הלייבלים של הקלאס ה i ל-1 ואת השאר למינוס 1
 - הרץ SVM PREDICT ושמו רק את הערך ההסתברותי גודל X^2 מספר הטסטים.
 - קח את העמודה החיובית (שהטסט נמצא בקלאס הנוכחי).
- עבור על כל הטסטים:
 - מצא את המקסימום בשורה ה i במטריצת ההסתברויות.
 - האינדקס של המקסימום הוא הקלאס שחזינו עבור הטסט ה i
 - אם האינדקס שווה לערך שנמצא ב $Y(i)$, אז חזינו נכון.
- קיבלנו 79%

```
Test accuracy is 0.79
```

ג. `:trainRBFsVM.m`

- הרצנו CV לכל קלאס עם 6 אפשרויות (כל C עם כל G).
 - מיינו את ה-DATA כאשר הקלאס הנוכחי היה בראש ה-DATA הממויין (סמפלים ולייבלים בהתאמה).
 - הרצנו את SVMTRAIN עם כל ערכי ה-C וערכי הגאמה
 - עבור $C=5$ וגאמה=0.1 קיבלנו `cv_acc = 97.18`
- אימנו 10 מסווגים ליניאריים עם $C=5$ וגאמה=0.1 ושמרנו אותם.

```
***
RBF SVM was trained with bestC = 5, bestG = 0.010
***
Cross-validation accuracy of RBF SVM is 97.18
```

ד. `m` (על ג):

- אותו הסבר כמו ב בדיוק חוץ מהתוצאה כמובן.
- קיבלנו 84%

```
Test accuracy of RBF SVM is 0.84
ALL DONE
```